Nano Research

Open Access

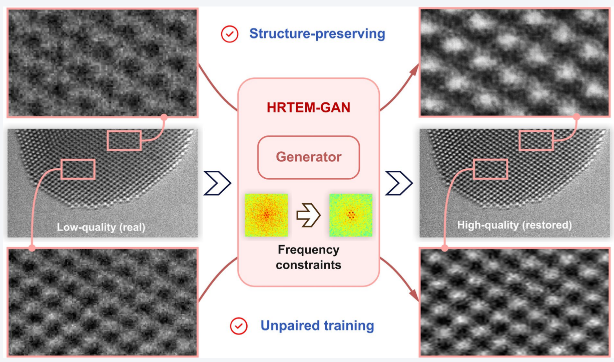
Research Article

HRTEM-GAN: Structure-preserving restoration of low-quality atomic-scale HRTEM images

Changtai Li[1,2,3,§], Yifan Zhang[3,§], Yu Jia[1,4], Yu Guo[1], Chao Yao[2] ✉, Xiaojuan Ban[1], and Yang He[1,4] ✉

[1] *Beijing Advanced Innovation Center for Materials Genome Engineering, School of Advanced Materials Innovation, University of Science and Technology Beijing, Beijing 100083, China*
[2] *School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China*
[3] *School of Intelligence Science and Technology, University of Science and Technology Beijing, Beijing 100083, China*
[4] *Department of Materials Science, School of Materials Science and Engineering, University of Science and Technology Beijing, Beijing 100083, China*
[§] *Changtai Li and Yifan Zhang contributed equally to this work.*

Cite this article: *Nano Research*, **2026**, *19*, 94908715. https://doi.org/10.26599/NR.2026.94908715



ABSTRACT: High-resolution transmission electron microscopy (HRTEM), with its outstanding spatial and temporal resolution, has enabled unprecedented understanding of the atomic structures of materials and how structure relates to properties and functions. However, capturing dynamic processes with high temporal resolution inevitably leads to severely degraded HRTEM images due to experimental conditions and imaging parameters, which substantially limit accurate structural analysis. In this paper, we develop a structure-preserving HRTEM restoration framework that enhances low-quality HRTEM images with blurred or incomplete atomic arrangements using a generative deep learning approach while maintaining physical fidelity. Specifically, we propose HRTEM-generative adversarial network (GAN), a cycle-consistent generative framework that operates under unpaired training conditions and performs patch-level distribution modeling between high- and low-quality image domains, while explicitly incorporating frequency-domain constraints to preserve atomic-scale structural fidelity. This design enables effective restoration of low-quality HRTEM images, yielding structurally coherent atomic lattices that are fully suitable for subsequent recognition and quantitative analysis. The proposed method is validated on a real experimental dataset acquired during *in situ* imaging of Au catalysts under CO oxidation conditions. Compared with representative methods, HRTEM-GAN achieves substantial improvements in image restoration quality and consistently enhances downstream atomic column recognition performance. These results demonstrate the potential of the proposed framework to facilitate reliable atomic-scale analysis in HRTEM studies.

KEYWORDS: high-resolution transmission electron microscopy (HRTEM), atomic-resolution image restoration, unpaired image-to-image translation, generative adversarial networks, artificial intelligence

1 Introduction

Understanding and controlling material properties through atomic-scale structure has long been a central pursuit in materials science, as atomic arrangements fundamentally govern material functionality and performance [1–4]. High-resolution transmission electron microscopy (HRTEM) possesses high spatial and atomic-scale resolution, therefore, plays a pivotal role in reliable materials characterization [5–7]; however, realistic experimental constraints often result in severely degraded images, which in turn hinder accurate identification of atomic sites and subsequent structural analysis [4, 8]. For instance, in catalytic systems, chemical reactions are often accompanied by rapid atomic-scale structural rearrangements occurring on timescales of tens of milliseconds [9–11]. Capturing image sequences at such high temporal resolution inevitably yields images that are strongly affected by shot noise, further complicating atomic-scale feature extraction and statistical analysis. It is generally impractical to suppress noise by

Received: March 7, 2026; **Revised:** April 3, 2026
Accepted: April 6, 2026

✉Address correspondence to Chao Yao, yaochao@ustb.edu.cn; Yang He, yanghe@ustb.edu.cn

increasing the incident electron beam intensity, as high-dose exposure to high-energy electrons can induce damage to the material. This motivates the need to recover high-quality (HQ) atomic-resolution information from degraded low-quality (LQ) HRTEM images while preserving physical fidelity [12, 13].

In recent years, a variety of computational methods, including conventional image processing algorithms [14–16] and deep learning techniques [17–19], have been applied to mitigate the inherent limitations of microscopy imaging and improve image quality. Conventional rule-based methods, such as Wiener filtering, bilateral filtering, and block matching and three-dimensional filtering (BM3D), have been widely used for noise suppression and basic image enhancement [16, 20]. While effective in attenuating random noise, these methods rely on fixed assumptions and limited local statistics, making it difficult to robustly handle complex backgrounds and severe image degradation. Recently, deep learning (DL)-based methods, leveraging their strong non-linear modeling capacity, have been extensively explored for electron microscopy image processing, primarily targeting low-level image enhancement tasks such as denoising, super-resolution, and generation [17].

Representative DL-based denoising approaches for TEM typically rely on supervised learning with synthetic or simulated training data [18, 19, 21]. For example, Lin et al. [18] introduced the AtomSegNet framework, in which convolutional encoder–decoder networks are trained on physics-informed simulated images. Mohan et al. [12] further proposed a simulation-based denoising (SBD) framework, demonstrating that CNNs trained on carefully designed forward models can effectively denoise low-SNR TEM images and generalize to real experimental data, particularly when large receptive fields are used to capture non-local atomic periodicities. More recently, frequency-aware enhancement strategies have been explored, incorporating spatial-frequency interactions to better exploit the periodic nature of atomic arrangements. Li et al. [22] proposed a framework that integrates a spatial-frequency interaction network with noise calibration-based data synthesis to model frequency-domain information. Nevertheless, most existing methods still struggle under severely degraded imaging conditions. Importantly, most methods rely on synthetic images to construct paired samples for supervision, while overlooking the rich and realistic texture characteristics present in real high-quality experimental images.

To alleviate the dependence on paired supervision, unpaired image-to-image translation frameworks have been actively studied [23–26]. In this context, approaches based on the Cycle-consistent generative adversarial network (CycleGAN) [26–28] have attracted increasing attention as a pioneering framework capable of translating images between two domains without requiring paired training samples. Building upon this paradigm, Quan et al. [29] proposed an asymmetrically cyclic adversarial network for unpaired denoising in electron microscopy, which extends the CycleGAN framework to learn forward and inverse noise mappings for artifact suppression without paired clean targets. However, the method still relies primarily on cycle-consistency for structural preservation and lacks explicit atomic-scale geometric or physical constraints; their whole-image-level modeling with weak implicit constraints limits controllability at the atomic scale, frequently resulting in inaccurate atomic reconstruction and inconsistent atomic positions.

In this paper, we propose HRTEM-GAN, a CycleGAN-based framework designed for structure-preserved image restoration in atomic-resolution HRTEM imaging. Instead of performing whole-image translation, HRTEM-GAN adopts patch-level modeling, which reduces the effective receptive scope of generation and allows the network to focus on fine-grained atomic structures, alleviating incomplete atomic recovery commonly observed in full-image generation. To further improve the modeling of long-range atomic correlations, a Vision Transformer (ViT) [30] module is incorporated at the bottleneck of the generator, enabling global dependency modeling beyond local convolutional neighborhoods. In addition, we explicitly introduce frequency-domain modeling and constraints to exploit the intrinsic spectral characteristics of HRTEM images. By enforcing spatial-frequency consistency via a dedicated frequency branch and frequency-aware losses, the proposed framework balances image quality enhancement and atomic-scale structural fidelity. We have performed extensive quantitative and qualitative evaluations on a real experimental dataset, demonstrating that HRTEM-GAN outperforms existing methods of TEM image quality enhancement and atomic structure preservation. Evaluations using AtomSegNet further show improved atomic recognition accuracy and stability, underscoring the value of the proposed framework for reliable quantitative atomic-scale analysis.

2 Experimental

2.1 Pipeline overview and dataset preparation

Our framework comprised two stages arranged in a sequential pipeline (see Fig. 1). Firstly, the acquired HRTEM video frames were automatically divided into LQ and HQ groups (unpaired) and cropped into patches that were then classified into foreground (nanocrystal area) and background (support area). Further details of this preprocessing were provided in Notes S1 and S2 in the Electronic Supplementary Material (ESM). Secondly, the proposed HRTEM-GAN approach was trained exclusively on unpaired LQ and HQ foreground patches, enabling the network to focus on learning atomic patterns. We employed CycleGAN to establish bidirectional mappings between the two domains, using two generators and two discriminators constrained by cycle consistency. We next detailed our HRTEM-GAN (Section 2.2), frequency-enhanced feature interaction (Section 2.3), and implementation details (Section 2.4).

Our dataset was obtained from *in situ* aberration-corrected TEM experiments, capturing surface step-site dynamics of Au catalysts during CO oxidation at room temperature. *In situ* TEM observations were performed on a Titan ETEM G2 microscope equipped with an imaging-side spherical aberration corrector, and image sequences were recorded using a charge-coupled device (CCD) camera at a frame rate of 10 frames per second, corresponding to an effective exposure time of approximately 0.1 s per frame. During the experiment, a nanosized Au crystal with a [011] orientation was manipulated to contact the mobile Au probe, forming bi-crystals through Joule heating inside the TEM chamber under a vacuum of approximately 10^{-6} mbar. Due to the differences in imaging parameters, the obtained HRTEM images exhibited significant quality variations. The images were therefore screened based on spatial resolution, atomic-column clarity, and lattice integrity, and subsequently divided into HQ and LQ subsets that were intrinsically unpaired and non-aligned, forming an unpaired training dataset.

2.2 HRTEM-GAN

To address the challenge of obtaining perfectly aligned pairs of

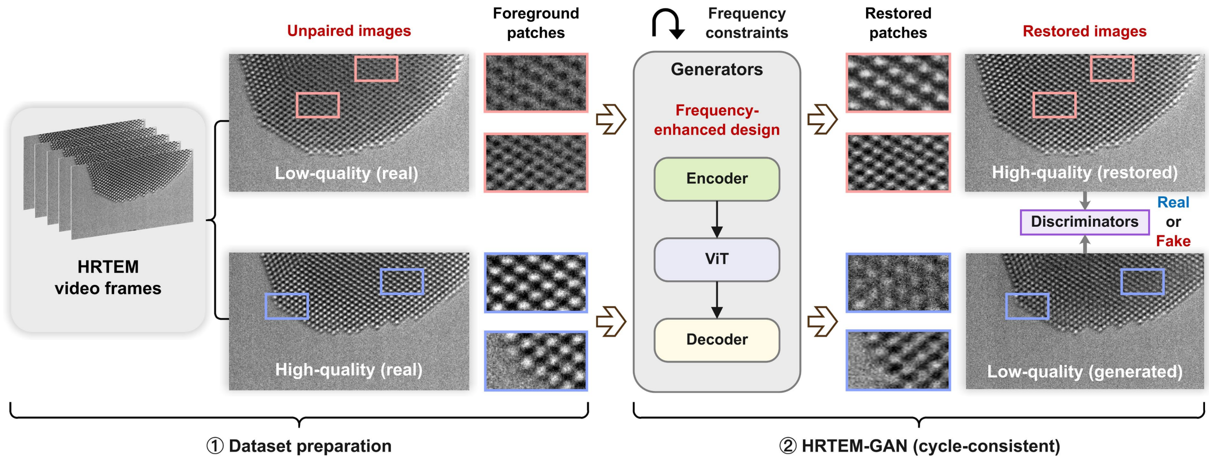


Figure 1 Overall pipeline. Unpaired LQ and HQ foreground patches from HRTEM video frames are used to train HRTEM-GAN via cycle-consistent learning, jointly optimizing bidirectional LQ ↔ HQ generators (Encoder–ViT–Decoder) with discriminators under adversarial and frequency-domain constraints to restore images. The proposed HRTEM-GAN is further detailed in Fig. 2.

noisy experimental images and clean ground-truth images in HRTEM experiments—which renders standard supervised learning methods (e.g., Pix2Pix [31]) inapplicable—we propose HRTEM-GAN (see Fig. 2(a)). It is based on CycleGAN and learns robust cross-domain mappings using unpaired data. In particular, the cycle-consistency constraint encouraged an image translated from the source domain to the target domain and back to remain faithful to the input, which is essential for preserving the geometric arrangement of atomic columns during enhancement. Here, we consider unpaired translation between the low-quality and high-quality HRTEM domains, denoted as LQ and HQ, respectively. We trained two generators $G_{\mathrm{H}} \equiv G_{\mathrm{LQ \to HQ}}$ and $G_{\mathrm{L}} \equiv G_{\mathrm{HQ \to LQ}}$, together with discriminators D_{H} and D_{L}. The base objective is formulated as

$$L_{\mathrm{base}} = L_{\mathrm{GAN}}(G_{\mathrm{H}}, D_{\mathrm{H}}) + L_{\mathrm{GAN}}(G_{\mathrm{L}}, D_{\mathrm{L}}) + \lambda_{\mathrm{cyc}} L_{\mathrm{cyc}}(G_{\mathrm{H}}, G_{\mathrm{L}}) + \lambda_{\mathrm{idt}} L_{\mathrm{idt}}(G_{\mathrm{H}}, G_{\mathrm{L}}) \tag{1}$$

where L_{GAN}, L_{cyc}, and L_{idt} are defined in Note S3 in the ESM.

Nevertheless, standard CycleGAN architectures often struggled to fully resolve the high-frequency periodic patterns intrinsic to atomic lattices. To address this limitation, HRTEM-GAN introduced two major modifications.

Firstly, we imposed feature-level alignment on the fused bottleneck representations of the two generators G_{H} and G_{L} to encourage the generated samples to approximate the feature characteristics of the target domain. z_{H} and z_{L} denote the corresponding bottleneck features. Feature consistency was enforced from two complementary aspects: frequency-domain alignment and distributional alignment.

Frequency-domain alignment. We applied the fast Fourier transform (FFT) to the bottleneck features and penalize the discrepancy between their magnitude spectrum, thereby encouraging the translated results to preserve consistent spectral statistics and reduce unrealistic high-frequency artifacts. Formally, $\mathcal{F}(\cdot)$ denotes the FFT and $\cdot$ its magnitude spectrum. The frequency-domain alignment loss is defined as

$$L_{\mathrm{FFT}} = E_{z_{\mathrm{H}}, z_{\mathrm{L}}}\left[\left\| |F(z_{\mathrm{H}}(x))| - |F(z_{\mathrm{L}}(y))| \right\|_1\right] \tag{2}$$

where $\|\cdot\|_1$ denotes the element-wise l_1 norm.

Distributional alignment via the characteristic function (CF). In addition, we introduced the CF [32] as a robust measure to match feature-space distributions between the two domains. The CF yielded complex-valued responses whose real and imaginary parts encode cosine and sine components, naturally supporting an amplitude–phase decomposition. The amplitude characterized how strongly and broadly the feature distribution responds to each probing direction, which correlates with the coverage of fine-grained textures and local contrast variations. The phase, by contrast, was more sensitive to structural shifts and thus penalizes misalignment that would manifest as lattice distortion or atomic-position drift. Accordingly, we enforced joint amplitude and phase consistency between the bottleneck features of G_{H} and G_{L}. This motivated our design to enforce joint amplitude and phase consistency between the two domains

$$L_{\mathrm{CF}} = \left\| |\mathrm{CF}(z_{\mathrm{H}})| - |\mathrm{CF}(z_{\mathrm{L}})| \right\|_1 + \lambda_{\phi}\left(1 - \cos\left(\angle \mathrm{CF}(z_{\mathrm{H}}) - \angle \mathrm{CF}(z_{\mathrm{L}})\right)\right) \tag{3}$$

where $|\cdot|$ and $\angle(\cdot)$ denote the magnitude and phase of the complex-valued CF output, respectively, and λ_{ϕ} balances the phase term. Details of the CF construction are deferred to Note S4 in the ESM.

In summary, the total loss of our training is

$$L_{\mathrm{total}} = L_{\mathrm{base}} + \lambda_{\mathrm{CF}} L_{\mathrm{CF}} + \lambda_{\mathrm{FFT}} L_{\mathrm{FFT}} \tag{4}$$

where λ_{CF} and λ_{FFT} are weighting hyperparameters that balance the contributions of the CF constraint and the frequency-domain consistency constraint to the overall training objective.

Secondly, our HRTEM-GAN moved beyond CycleGAN whose conventional U-Net [33] generator by integrating a dual-branch architecture in each encoder layer (see Fig. 2(c)). Considering that spectral features naturally encode global periodicity, we introduced an independent frequency branch alongside the spatial branch to preserve the intrinsic long-range order and fine-grained details of crystal lattices (detailed in the following section). To further enhance the generator's ability to recover lattice-level structural regularity, a ViT [30] module was incorporated at the bottleneck of the dual-branch encoder–decoder architecture. While convolutional operations effectively captured local atomic textures, their limited receptive field constrained the modeling of long-range periodicity. Addressing this, the fused spatial-frequency features at the bottleneck were flattened into a token sequence, and two-dimensional positional encodings were added to preserve the

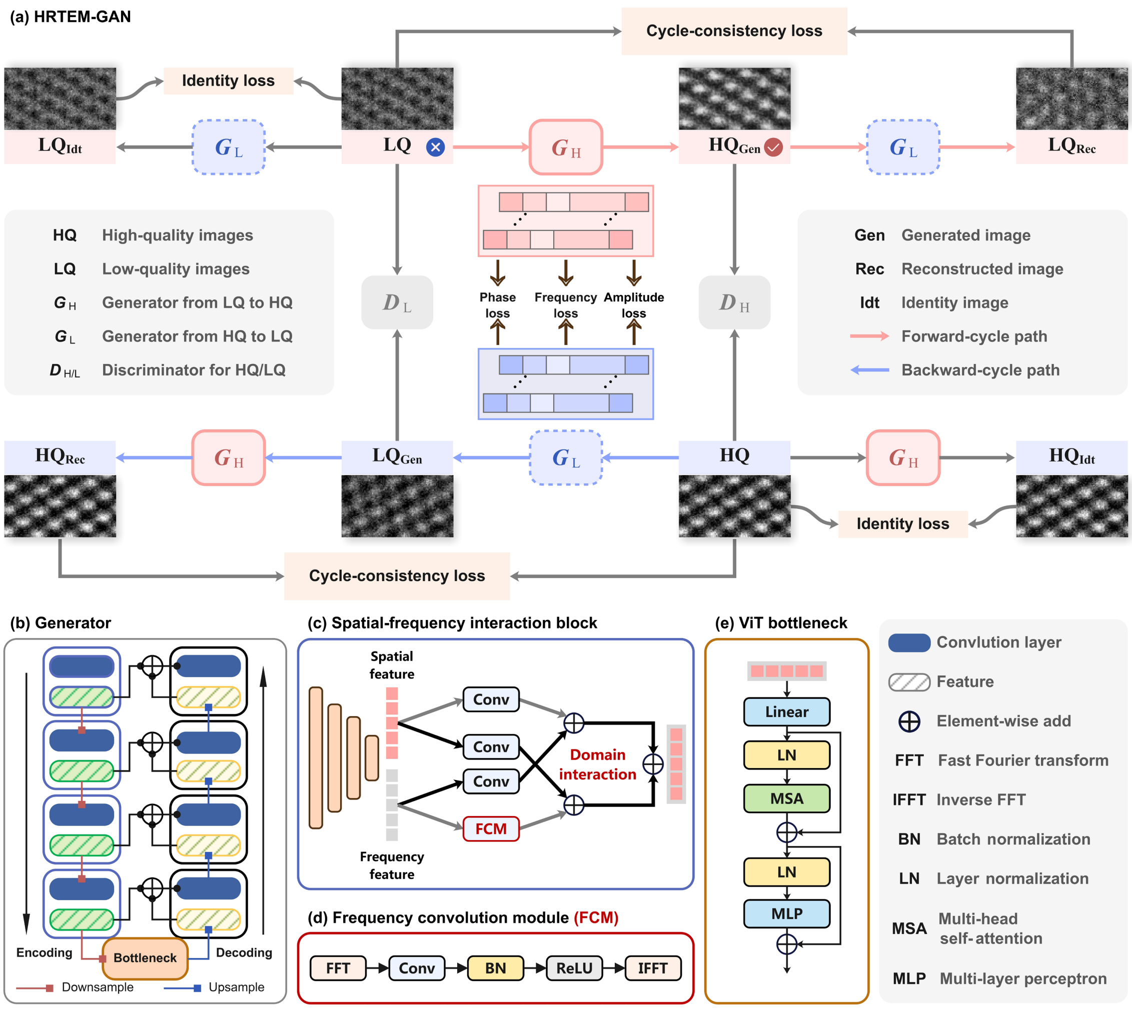


Figure 2 Network architecture of the proposed HRTEM-GAN and its key components. (a) Overview of the HRTEM-GAN framework. (b) Enhanced UNet-ViT generator. (c) Spatial-frequency interaction block. (d) FCM. (e) ViT bottleneck.

original lattice geometry. The token sequence was then processed by several stacked multi-head self-attention (MSA) layers, enabling the network to learn global dependencies between arbitrary spatial locations and to represent crystal lattices with coherent long-range ordering. The globally contextualized features were subsequently reshaped back into a two-dimensional feature map and propagated into the decoder through skip connections. By embedding global self-attention at the structural bottleneck, this hybrid design enforced global lattice consistency while retaining high-fidelity local atomic details provided by the convolutional and frequency branches.

2.3 Frequency-enhanced feature interaction

Complementing the independent frequency branch in each layer of the encoder, the input HRTEM image was processed by a standard convolutional layer to obtain primary features, which are then split along the channel dimension to initialize parallel spatial and frequency streams. After that, we employed the spatial-frequency interaction block [22] to progressively enhance both feature streams. As illustrated in Fig. 2(d), convolutional modules are used to capture local geometric structures and atomic neighborhoods, while frequency convolution modules (FCM) emphasize periodic lattice patterns and high-frequency structural details by leveraging Fourier transform.

To prevent semantic drift and maintain mutual consistency, each interaction block incorporated a cross-branch information exchange mechanism. Specifically, features from one branch were passed through an additional convolutional layer and added to the other branch, enabling reciprocal refinement. This allowed spatial features to benefit from global periodic cues, while frequency-domain features remain guided by local structural geometry.

After the interaction stages, the enhanced spatial-frequency features were fused into a unified representation and propagated through the decoder to reconstruct the output image. This integrated design significantly improved the generator's capacity to restore atomic lattice periodicity, preserve sharp high-frequency details, and maintain local structural fidelity under unpaired training conditions. The restored image patches were finally reassembled into full-resolution frames, yielding high-quality HRTEM images with globally coherent and physically meaningful atomic arrangements.

2.4 Implementation details

Our model was trained on unpaired data. All experiments were implemented in PyTorch and trained on an NVIDIA 4090 GPU with CUDA acceleration. The networks were optimized using the Adam optimizer with an initial learning rate of 2×10^{-4} and $\beta_1 = 0.5$, where the learning rate was kept constant for the first

150 epochs and then linearly decayed to zero over the following 20 epochs. An unaligned dataset setting with a high-to-low translation direction was adopted, and all input images were resized to 286×286 and randomly cropped to 256×256 during training, with a batch size of 128. The model followed the standard CycleGAN configuration with 64 base feature channels for both generator and discriminator, employed a PatchGAN discriminator, and was trained using the least-squares GAN objective, together with an image buffer of size 50 to stabilize training. The weighting hyperparameters in Eqs. (1) and (4) were empirically set to $\lambda_{\text{cyc}} = 10$, $\lambda_{\text{idt}} = 5$, $\lambda_{\text{CF}} = 0.5$, and $\lambda_{\text{FFT}} = 0.5$. Note that although HRTEM-GAN is trained with bidirectional mappings between LQ and HQ domains, practical deployment primarily uses the LQ → HQ branch.

3 Results and discussion

3.1 Evaluation protocols

To comprehensively evaluate the proposed method, we conducted both image generation quality assessment and recognition performance evaluation. For generation, quantitative metrics were adopted to measure the distribution consistency between generated images and real images, complemented by qualitative feature-space visualization. Specifically, Fréchet Inception Distance (FID) [34] and Kullback–Leibler (KL) divergence [35] were used for quantitative evaluation, while t-SNE visualization [36] was employed for qualitative analysis. For recognition, standard pixel-level metrics were adopted to assess the accuracy between predicted masks and ground-truth annotations, including Precision, Dice coefficient, and Intersection over Union (IoU). The formulas for computing the evaluation metrics are provided in Note S5 in the ESM.

As comparison baselines, we compared HRTEM-GAN with both conventional denoising algorithms and learning-based restoration methods. We included Wiener filtering and ABSF as representative classical frequency-/filtering-based approaches, and BM3D as a widely used patch-based denoiser. Besides, we evaluated three learning-based alternatives, i.e., AtomSegNet [18], SBD [12], and SFIN [22], which have been adopted in prior work for improving TEM image quality. To assess recognition performance in a controlled manner, we used the same segmentation network (AtomSegNet) to predict atomic masks from the restored images produced by each method, and then computed pixel-level metrics (Precision, Dice, and IoU) against the ground-truth.

3.2 Quantitative results of restoration

We first quantify the distribution-level fidelity of restored images to the HQ domain using feature-based metrics, aiming to evaluate whether a restoration method genuinely narrows the gap between LQ inputs and HQ references in the learned representation space. Specifically, we report the FID and KL divergence computed between restored foreground patches and real HQ patches. The quantitative comparison across different methods is summarized in Table 1.

Among the conventional denoising methods, BM3D achieves the strongest performance, while Wiener and ABSF exhibit substantially larger distribution gaps (FID > 215 and KL $\geq$ 1.320). The learning-based counterparts (AtomSegNet, SBD, and SFIN) reduce KL divergence to a moderate level but still yield relatively high FID values, indicating that the restored outputs remain far from the target HQ distribution in feature space. In contrast, HRTEM-GAN achieves the lowest FID (33.622) and KL divergence (0.092), outperforming all competing approaches by a large margin and demonstrating substantially improved distribution alignment and restoration fidelity.

Table 1 also reports the average inference time per image for each method. HRTEM-GAN requires 2.928 s per image, which is comparable to SFIN and SBD, and faster than AtomSegNet. Although BM3D achieves relatively good restoration among the conventional methods, its inference time is much longer (32.491 s per image). Hence, despite incorporating a ViT module, HRTEM-GAN still maintains a practical inference speed, while achieving the best restoration performance among all compared methods.

3.3 Qualitative results of restoration

3.3.1 Patch-level qualitative comparison

We first inspect atomic-level detail recovery through local magnification. As shown in Fig. 3, two regions are cropped from each raw HRTEM image restored by different methods. It is observed that BM3D and other deep learning-based methods are effective in noise reduction but fail to recover the underlying atomic structures.

Two representative cases further highlight the difference. In region (3), where the raw observation is severely blurred and atomic contrast is barely discernible, other methods mainly suppress noise but still fail to reveal stable lattice periodicity. In contrast, our HRTEM-GAN recovers clearer atomic-point contrast with sharper boundaries, making the underlying lattice pattern more distinguishable. In region (6), adjacent atomic columns appear partially merged due to imaging degradations; other competing methods tend to preserve the adhesion after denoising, whereas our HRTEM-GAN better separates the fused atomic points and restores more consistent inter-column boundaries. Similar trends are observed across the remaining regions, showcasing the overall robustness of HRTEM-GAN in restoring both atomic-level contrast and physically consistent lattice periodicity under diverse degradation patterns.

3.3.2 Image-level qualitative comparison

To evaluate overall restoration quality beyond local patches, we compare full-image reconstructions produced by HRTEM-GAN and other approaches. Restored patches are aggregated into a complete image using an overlapping sliding-window strategy with pixel-wise averaging in overlapping regions. After removing padded

Table 1 Quantitative comparison of restoration performance and inference time per image between our method and existing approaches

Type	Method	FID ↓	KL[a] ↓	Time (s)
Conventional	Wiener	215.825	1.530	3.792
	ABSF	223.188	1.320	3.821
	BM3D	<u>132.504</u>[c]	0.424	32.491
Learning-based	AtomSegNet	172.224	<u>0.326</u>[c]	6.154
	SBD	169.088	0.402	3.016
	SFIN	211.103	0.348	**2.640**[b]
	Ours	**33.622**[b]	**0.092**[b]	<u>2.928</u>[c]

[a]denotes KL divergence. [b]The best results. [c]The second-best results.

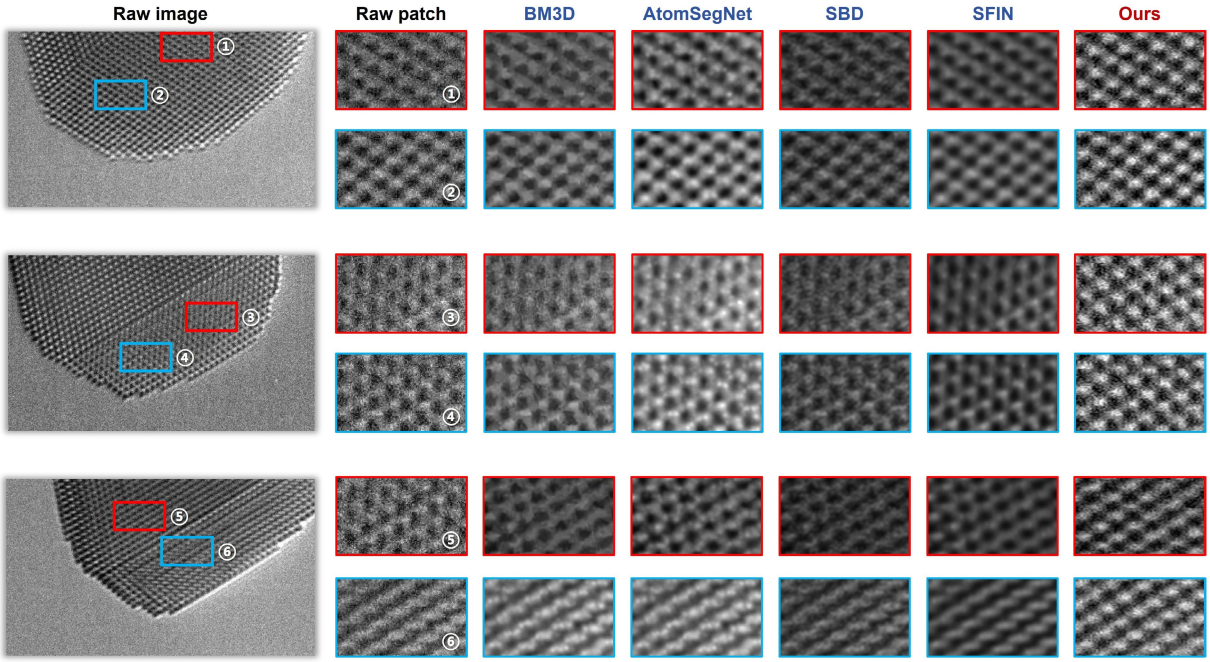


Figure 3 Qualitative comparison of patch-level restoration performance between our method and existing approaches. The red and blue bounding boxes indicate two selected patches for each example.

boundaries, the reconstructed image is resized to the original resolution to obtain the final output.

Figure 4 presents representative image-level results across four test images. It can be observed that BM3D and the learning-based methods (AtomSegNet, SBD, and SFIN) reduce noise to some extent, but their outputs remain largely smoothing-dominated, with attenuated atomic contrast and locally unstable lattice textures. This limitation is particularly evident in Image 2 and Image 3, where the highlighted regions emphasize challenging structures. In Image 2 (red box), the competing methods tend to over-smooth the lattice fringes, leading to weakened peak-valley contrast and less well-defined atomic columns in the interior lattice region. In contrast, HRTEM-GAN preserves clearer atomic-point contrast and more regular lattice periodicity within the boxed area, maintaining sharper boundaries between adjacent columns while suppressing background noise. In Image 3 (blue box), the specimen exhibits a twin-boundary-like (or stacking-fault-like) interfacial contrast, where the lattice orientation and fringe continuity change across a narrow region. Such defect- or interface-related patterns are particularly challenging under low SNR, as they require preserving not only periodic lattice fringes but also the coherence across the interface. The other methods tend to over-smooth this area and partially wash out the interfacial contrast, resulting in blurred textures and locally disrupted periodicity. Oppositely, HRTEM-GAN better maintains the interface-induced contrast variation while preserving the surrounding lattice order, yielding a more coherent reconstruction with clearer atomic features in the highlighted region. Overall, these observations indicate that the proposed method goes beyond generic denoising and yields restorations that are visually closer to high-quality TEM observations.

3.3.3 *Spatial-frequency consistency analysis*

Beyond spatial-domain appearance, a key indicator of physically meaningful HRTEM restoration is whether the method recovers the characteristic lattice periodicity in the frequency domain. Therefore, we compute the two-dimensional (2D) FFT log-magnitude after applying a two-dimensional Hann window and visualize it with the zero-frequency component centered. This process suppresses boundary-induced leakage and yields more stable reciprocal-space patterns for assessing lattice-related high-frequency content.

Under this spatial-frequency inspection, as shown in Fig. 5, the real LQ experimental patches (the first row) exhibit attenuated and diffuse high-frequency responses, where reciprocal-lattice signatures are weak and partially blurred, consistent with diminished fringe contrast in the spatial domain. In contrast, the outputs restored by G_H (the second row) show clearer atomic-column modulation and sharper lattice fringes, accompanied by more discernible Bragg-like peaks (or reciprocal-lattice spots) and reduced spectral smearing in the log-magnitude representation. Compared with the degraded input, the restored images exhibit a more regular periodic contrast pattern in real space and a more concentrated distribution of lattice-related frequency components in reciprocal space. These observations suggest that our HRTEM-GAN improves image quality by reconstructing lattice periodicity rather than merely amplifying contrast.

From a physical standpoint, genuine lattice restoration should manifest as localized energy re-concentration at lattice-consistent spatial frequencies, rather than an uniform amplification of high-frequency components. Accordingly, compared with LQ inputs, the restored results exhibit not only increased peak prominence but also noticeably reduced peak broadening and spectral smearing, indicating a more coherent periodic modulation in the spatial domain. Moreover, the reciprocal-space patterns remain largely centro-symmetric and do not show systematic spurious peaks or ring-like over-enhancement, which suggests that the model tends to reconstruct plausible lattice periodicity instead of introducing artificial textures. This behavior is closely related to the stabilization of periodic lattice parameters during restoration. For crystalline Au

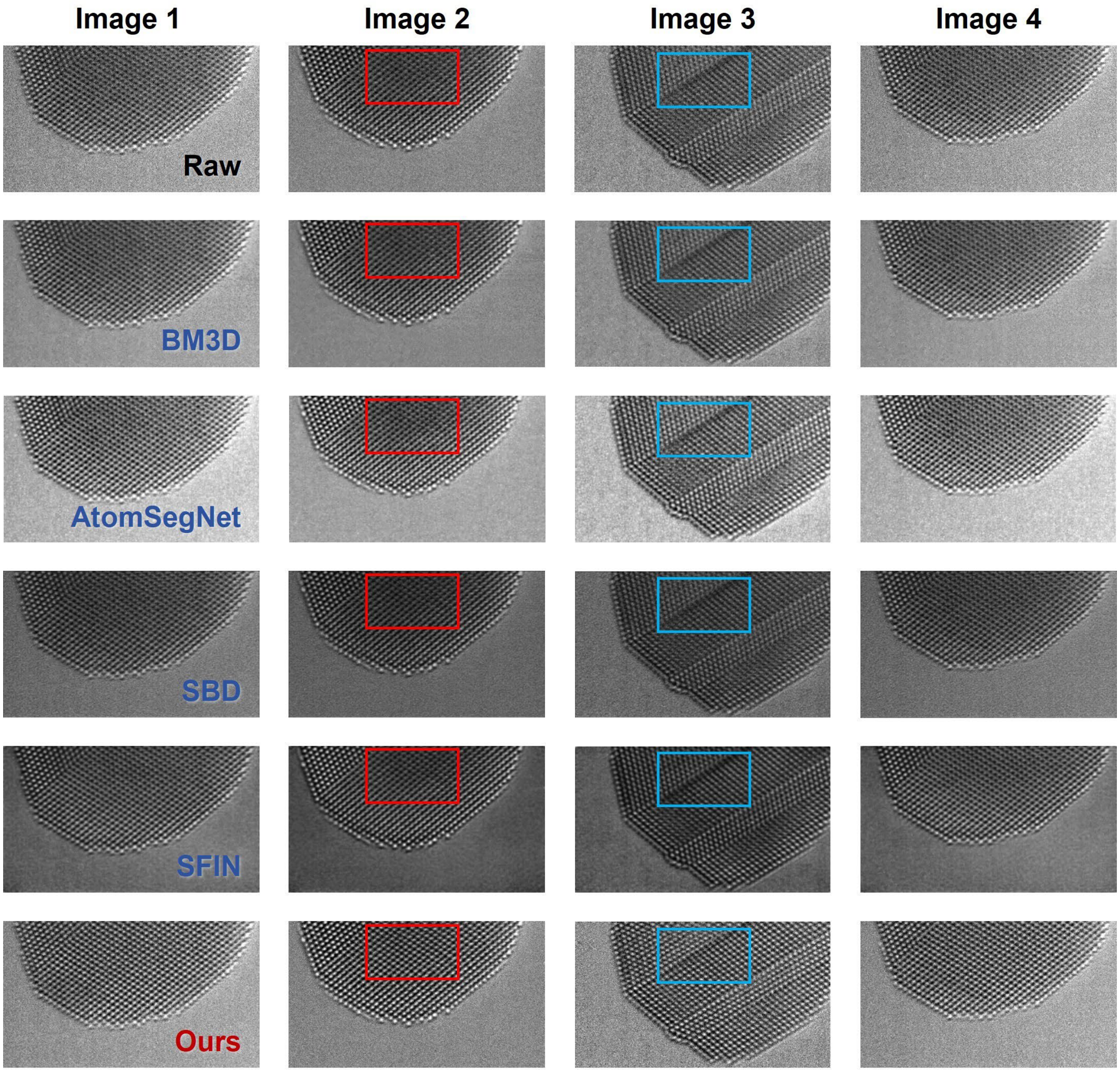


Figure 4 Qualitative comparison of image-level restoration performance between our method and existing approaches. Columns correspond to different test images, and rows correspond to different restoration methods (method names are indicated on the left). The red and blue bounding boxes mark representative regions in Image 2 and Image 3.

lattices, the periodic spacing and ordering in the spatial domain are reflected by the location, sharpness, and symmetry of reciprocal-space peaks. By constraining the restored features to remain consistent with lattice-related frequency components, the proposed frequency-domain supervision helps suppress spectral broadening and spurious frequency responses, thereby reducing local distortions of lattice spacing and preserving more stable periodic atomic arrangements in the restored images.

Representative HQ real patches are additionally provided as a reference (not paired with the corresponding LQ inputs) to illustrate typical reciprocal-space patterns of high-quality images, which are qualitatively better matched by the restored results.

3.3.4 *t-SNE visualization analysis*

To examine whether restored images are not only visually improved but also feature-wise aligned with the HQ domain, we visualize the deep-feature embeddings of foreground patches using t-SNE in Fig. 6, where red circles denote real HQ patches and blue triangles denote the corresponding patches obtained after restoration.

After applying BM3D and the learning-based methods (AtomSegNet, SBD, and SFIN), the restored features remain noticeably detached from the HQ cluster and often become more scattered, suggesting that these methods mainly perform local denoising/enhancement without effectively mapping the low-quality inputs toward the HQ feature distribution. In contrast, the embeddings produced by HRTEM-GAN show substantially increased overlap with the HQ points and reduced inter-domain discrepancy, indicating that the proposed model better recovers high-frequency structural characteristics and overall statistical properties consistent with the HQ domain, thereby achieving more effective distribution-level alignment.

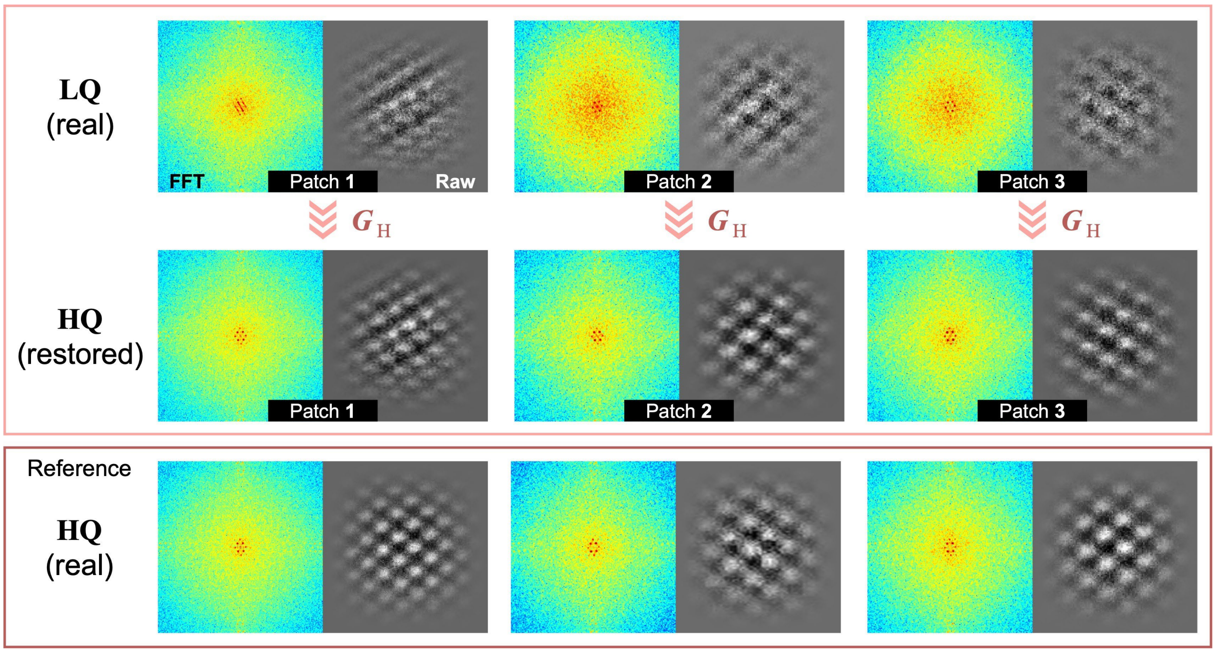


Figure 5 Spatial-frequency comparison of restoration results. The bottom row provides representative real HQ patches as a reference for the expected reciprocal-space patterns (unpaired with the LQ inputs).

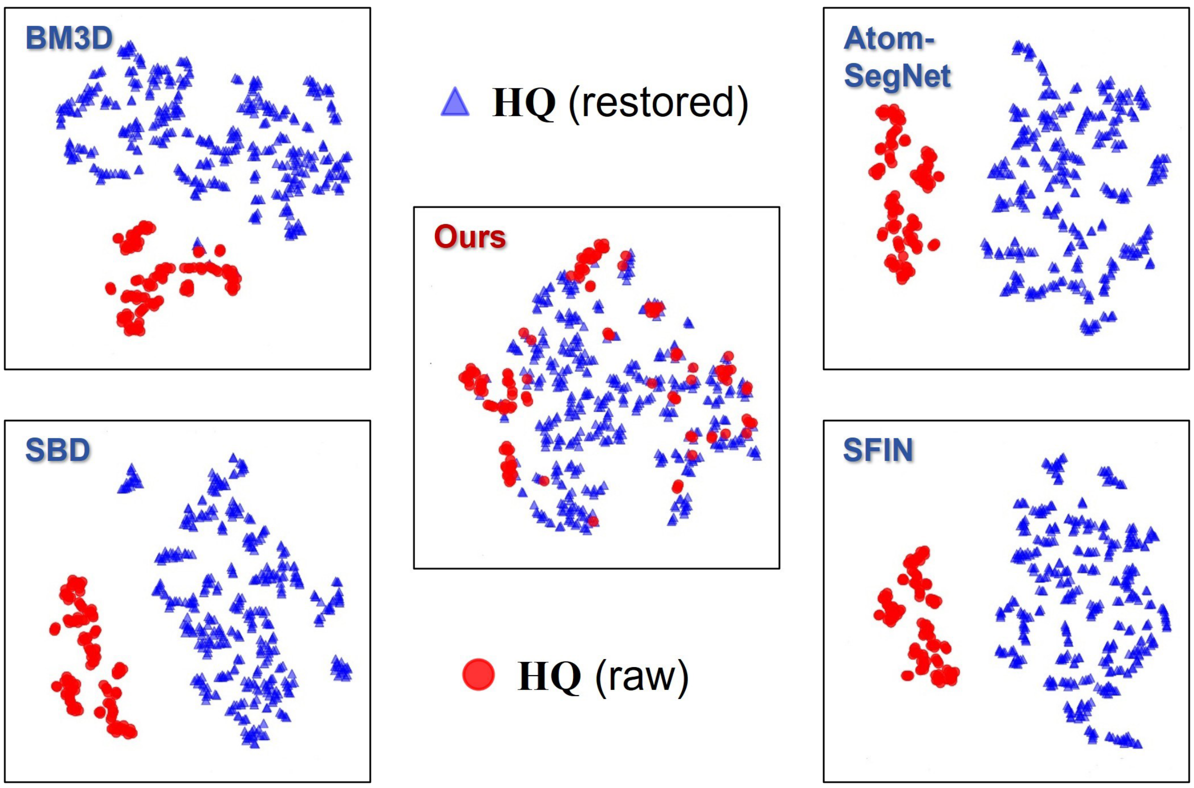


Figure 6 t-SNE visualization of deep-feature embeddings for atomic patches under different restoration methods. Red circles denote real HQ (raw) patches, and blue triangles denote the corresponding patches obtained after restoration.

3.4 Performance of atomic recognition

To further evaluate how well different restoration models preserve atomic-level structures, we apply the trained AtomSegNet to segment atomic columns on the restored outputs of representative DL method (SBD and SFIN) and quantify structural fidelity using segmentation-based metrics (Precision, Dice, and IoU) against the ground-truth labels, and additionally provide qualitative comparisons of the predicted masks to visually assess atomic-column completeness, separability, and boundary consistency under challenging imaging conditions.

Table 2 summarizes the segmentation accuracy on images restored by different methods. Compared with AtomSegNet, SBD, and SFIN, our method yields consistent gains across all metrics, indicating improved atomic-column structural fidelity after restoration. Specifically, our method achieves a Precision of 0.806, which is higher than SFIN (0.692) and AtomSegNet (0.599), suggesting fewer false positives and better separability between adjacent atomic columns. Likewise, the Dice score increases to 0.731 (vs. 0.691 for SFIN and 0.620 for AtomSegNet), and IoU improves to 0.578 (vs. 0.533 for SFIN and 0.455 for AtomSegNet), reflecting better mask overlap and more complete atomic-column recovery. These quantitative improvements are also consistent with the qualitative comparison below, where the restored images produced by HRTEM-GAN deliver more accurate and stable segmentation results.

Table 2 Quantitative comparison of segmentation performance under different restoration methods

Method	Precision	Dice ↑	IoU ↑
AtomSegNet	0.599	0.620	0.455
SBD	0.468	0.548	0.386
SFIN	0.692	0.691[b]	0.533[b]
Ours	**0.806**[a]	**0.731**[a]	**0.578**[a]

[a]The best results. [b]The second-best results.

As shown in Fig. 7, the atomic-column segmentation results of AtomSegNet, SBD, and SFIN exhibit typical failure modes under LQ conditions, including atomic-region merging, centroid shifts, and local disruption of lattice topology. These errors are not merely visual artifacts. The effective atomic-column count and nearest-neighbor statistics are distorted. In contrast, our method consistently produces more accurate and coherent segmentation results across three LQ patches. The segmentation binary maps

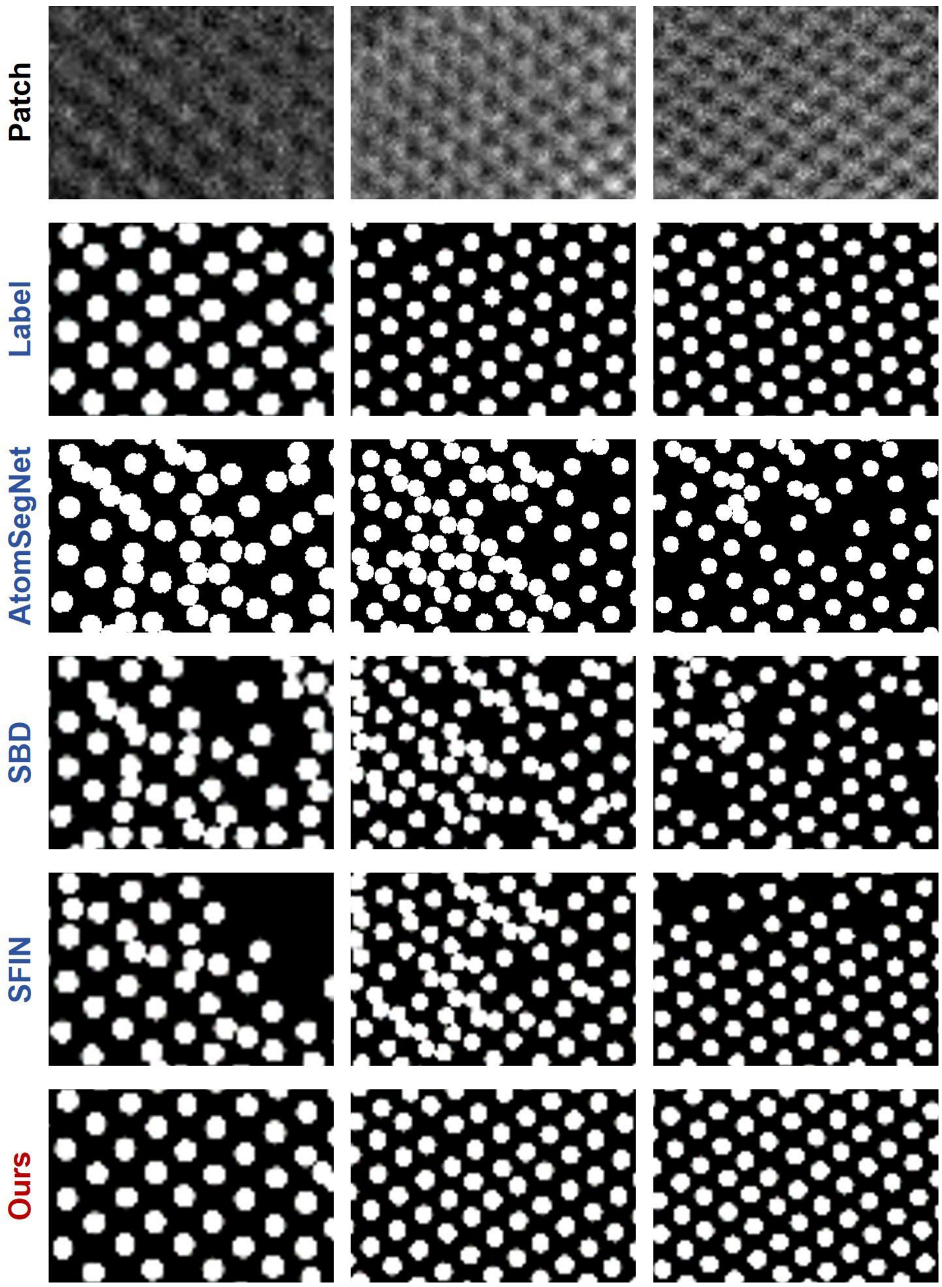


Figure 7 Qualitative comparison of atomic column segmentation performance.

produced on the restored images by our method exhibit uniformly distributed atomic regions and well-preserved lattice periodicity, maintaining complete and regular atomic structures.

Overall, the comparative segmentation results demonstrate that HRTEM-GAN outperforms other learning-based restoration methods. The images restored by HRTEM-GAN enable the atomic segmentation network to achieve more accurate and stable atomic identification, indicating that the proposed method better preserves structurally meaningful information in atomic-resolution images. This improved restoration quality not only benefits downstream atomic modeling and materials analysis but also provides a more reliable basis for kinetic measurements and structural evolution analysis. Since such analyses rely on the consistent identification of lattice features and their temporal changes, the enhanced image quality can help reduce ambiguity in frame-by-frame interpretation and support more robust analysis of dynamic processes.

4 Conclusions

In this paper, we have developed a HRTEM-image restoration method (HRTEM-GAN) built on the generative method, CycleGAN. Targeted at the atomic-resolution scenario, HRTEM-GAN restores low-quality HRTEM observations into high-quality counterparts while preserving critical structure information, without requiring strictly paired LQ/HQ training data. We combine spatial restoration with frequency-consistent supervision and align the restored features with real high-quality features to enhance the reality. The ViT-based design at the bottleneck helps capture long-range lattice regularity and improves structural faithfulness in complex regions. Extensive experiments demonstrate that the restored outputs exhibit improved visual fidelity and closer distributional alignment with real high-quality images in both pixel space and feature space. Although HRTEM-GAN does not directly increase the hardware-limited temporal resolution of TEM, it can improve the interpretability of frames acquired under short-exposure or high-frame-rate conditions, thereby potentially facilitating higher-temporal-resolution analysis in practice. Importantly, the enhanced image quality consistently improves downstream atomic-column segmentation performance, yielding higher Precision, Dice, and IoU compared with recent learning-based restoration baselines.

In future work, we will explore more condition-aware and physics-guided constraints [37, 38] (e.g., incorporating imaging parameter priors or forward models) as well as broader multi-domain training to improve robustness. Although the current study focuses on atomic-scale lattice images, the proposed framework may potentially be extended to low-contrast objects, such as adsorbed molecules, provided that sufficient signal is retained in the raw data. For such applications, their weaker contrast and less pronounced structural regularity may require more targeted training data and additional physics-informed constraints. We also plan to investigate lightweight deployment strategies to facilitate practical integration into routine HRTEM workflows. Overall, HRTEM-GAN provides a practical and effective pathway for restoring atomic-resolution HRTEM images and enhancing subsequent quantitative analysis.

Electronic Supplementary Material: Supplementary material (additional methodological details, training settings, and evaluation metrics) is available in the online version of this article at https://doi.org/10.26599/NR.2026.94908715.

Data availability

All data needed to support the conclusions in the paper are presented in the manuscript and the Electronic Supplementary Material. Additional data related to this paper may be requested from the corresponding author upon request.

Acknowledgements

This work is granted by the Advanced Materials-National Science and Technology Major Project of China (No. 2025ZD0620003) and the National Natural Science Foundation of China (Nos. 62332017 and 62303043). The computing work is supported by USTB MatCom of Beijing Advanced Innovation Center for Materials Genome Engineering.

Declaration of competing interest

All the contributing authors report no conflicts of interest in this work.

Author contribution statement

C. T. L.: Conceptualization, methodology, writing – review & editing. Y. F. Z.: Formal analysis, validation, visualization, writing – original draft. Y. J.: Data curation, investigation. Y. G.: Writing – review & editing. C. Y.: Project administration, writing – review & editing. X. J. B.: Supervision, writing – review & editing. Y. H.: Resources, supervision, writing – review & editing. All the authors have approved the final manuscript.

Use of AI statement

None.

References

[1] Sun, L. G.; Wu, G.; Wang, Q.; Lu, J. Nanostructural metallic materials: Structures and mechanical properties. *Mater. Today.* **2020**, *38*, 114–135.

[2] Liu, Y. L.; Niu, C.; Wang, Z.; Gan, Y.; Zhu, Y.; Sun, S. H.; Shen, T. Machine learning in materials genome initiative: A review. *J. Mater. Sci. Technol.* **2020**, *57*, 113–122.

[3] Morgan, D.; Jacobs, R. Opportunities and challenges for machine learning in materials science. *Annu. Rev. Mater. Res.* **2020**, *50*, 71–103.

[4] Liu, Y.; Zhao, T. L.; Ju, W. W.; Shi, S. Q. Materials discovery and design using machine learning. *J. Materiomics.* **2017**, *3*, 159–177.

[5] Liu, G. D.; Shih, A. J.; Deng, H. Q.; Ojha, K.; Chen, X. T.; Luo, M. C.; McCrum, I. T.; Koper, M. T. M.; Greeley, J.; Zeng, Z. H. Site-specific reactivity of stepped Pt surfaces driven by stress release. *Nature* **2024**, *626*, 1005–1010.

[6] Crozier, P. A.; Leibovich, M.; Haluai, P.; Tan, M.; Thomas, A. M.; Vincent, J.; Mohan, S.; Marcos Morales, A.; Kulkarni, S. A.; Matteson, D. S. et al. Visualizing nanoparticle surface dynamics and instabilities enabled by deep denoising. *Science* **2025**, *387*, 949–954.

[7] Ziletti, A.; Kumar, D.; Scheffler, M.; Ghiringhelli, L. M. Insightful classification of crystal structures using deep learning. *Nat. Commun.* **2018**, *9*, 2775.

[8] Ziatdinov, M.; Dyck, O.; Maksov, A.; Li, X. F.; Sang, X. H.; Xiao, K.;

Unocic, R. R.; Vasudevan, R.; Jesse, S.; Kalinin, S. V. Deep learning of atomically resolved scanning transmission electron microscopy images: Chemical identification and tracking local transformations. *ACS Nano* **2017**, *11*, 12742–12752.
[9] Ge, M.; Su, F.; Zhao, Z.; Su, D. Deep learning analysis on microscopic imaging in materials science. *Mater. Today Nano* **2020**, *11*, 100087.
[10] Yang, S. H.; Choi, W.; Cho, B. W.; Agyapong-Fordjour, F. O. T.; Park, S.; Yun, S. J.; Kim, H. J.; Han, Y. K.; Lee, Y. H.; Kim, K. K. et al. Deep learning-assisted quantification of atomic dopants and defects in 2D materials. *Adv. Sci.* **2021**, *8*, 2101099.
[11] Botifoll, M.; Pinto-Huguet, I.; Arbiol, J. Machine learning in electron microscopy for advanced nanocharacterization: Current developments, available tools and future outlook. *Nanoscale Horiz.* **2022**, *7*, 1427–1477.
[12] Mohan, S.; Manzorro, R.; Vincent, J. L.; Tang, B.; Sheth, D. Y.; Simoncelli, E. P.; Matteson, D. S.; Crozier, P. A.; Fernandez-Granda, C. Deep denoising for scientific discovery: A case study in electron microscopy. *IEEE Trans. Comput. Imaging* **2022**, *8*, 585–597.
[13] He, R. A.; Cheng, R.; Lyu, X. K.; Tan, W. M.; Yan, B. Efficient online training for zero-shot time-lapse microscopy denoising and super-resolution. In *Proceedings of the 39th AAAI Conference on Artificial Intelligence*, Philadelphia, USA, 2025, pp 3419–3427.
[14] Kilaas, R. Optimal and near-optimal filters in high-resolution electron microscopy. *J. Microsc.* **1998**, *190*, 45–51.
[15] Mevenkamp, N.; Binev, P.; Dahmen, W.; Voyles, P. M.; Yankovich, A. B.; Berkels, B. Poisson noise removal from high-resolution stem images based on periodic block matching. *Adv. Struct. Chem. Imaging* **2015**, *1*, 3.
[16] Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095.
[17] Choudhary, K.; DeCost, B.; Chen, C.; Jain, A.; Tavazza, F.; Cohn, R.; Park, C. W.; Choudhary, A.; Agrawal, A.; Billinge, S. J. L. et al. Recent advances and applications of deep learning methods in materials science. *npj Comput. Mater.* **2022**, *8*, 59.
[18] Lin, R. Q.; Zhang, R.; Wang, C. Y.; Yang, X. Q.; Xin, H. L. TEMImageNet training library and AtomSegNet deep-learning models for high-precision atom segmentation, localization, denoising, and deblurring of atomic-resolution images. *Sci. Rep.* **2021**, *11*, 5386.
[19] Lobato, I.; Friedrich, T.; Van Aert, S. Deep convolutional neural networks to restore single-shot electron microscopy images. *npj Comput. Mater.* **2024**, *10*, 10.
[20] Kushwaha, H. S.; Tanwar, S.; Rathore, K. S.; Srivastava, S. De-noising filters for TEM (transmission electron microscopy) image of nanomaterials. In *Proceedings of the 2012 Second International Conference on Advanced Computing & Communication Technologies*, Rohtak, India, 2012, pp 276–281.
[21] Kazimi, B.; Ruzaeva, K.; Sandfeld, S. Self-supervised learning with generative adversarial networks for electron microscopy. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, USA, 2024, pp 71–81.
[22] Li, H. S.; Wu, Z. Q.; Shao, R. W.; Zhang, T.; Fu, Y. Noise calibration and spatial-frequency interactive network for STEM image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, USA, 2025, pp 21287–21296.
[23] Lee, H. Y.; Tseng, H. Y.; Huang, J. B.; Singh, M.; Yang, M. H. Diverse image-to-image translation via disentangled representations. In *Proceedings of the 15th European Conference on Computer Vision (ECCV)*, Munich, Germany, 2018, pp 36–52.
[24] Huang, X.; Liu, M. Y.; Belongie, S.; Kautz, J. Multimodal unsupervised image-to-image translation. In *Proceedings of the 15th European Conference on Computer Vision (ECCV)*, Munich, Germany, 2018, pp 179–196.
[25] Nizan, O.; Tal, A. Breaking the cycle-colleagues are all you need. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 2020, pp 7857–7866.
[26] Zhao, Y. H.; Wu, R. H.; Dong, H. Unpaired image-to-image translation using adversarial consistency loss. In *Proceedings of the 16th European Conference on Computer Vision (ECCV)*, Glasgow, UK, 2020, pp 800–815.
[27] Khan, A.; Lee, C. H.; Huang, P. Y.; Clark, B. K. Leveraging generative adversarial networks to create realistic scanning transmission electron microscopy images. *npj Comput. Mater.* **2023**, *9*, 85.
[28] Eliasson, H.; Lothian, A.; Surin, I.; Mitchell, S.; Pérez-Ramírez, J.; Erni, R. Precise size determination of supported catalyst nanoparticles via generative AI and scanning transmission electron microscopy. *Small Methods* **2025**, *9*, 2401108.
[29] Quan, T. M.; Hildebrand, D. G. C.; Lee, K.; Thomas, L. A.; Kuan, A. T.; Lee, W. C. A.; Jeong, W. K. Removing imaging artifacts in electron microscopy using an asymmetrically cyclic adversarial network without paired training data. In *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, Seoul, Korea (South), 2019, pp 3804–3813.
[30] Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X. H.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S. et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *Proceedings of the 9th International Conference on Learning Representations*, OpenReview, 2021.
[31] Isola, P.; Zhu, J. Y.; Zhou, T.; Efros, A. A. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp 1125–1134.
[32] Wang, S. B.; Yang, Y. C.; Liu, Z. Y.; Sun, C. H.; Hu, X. M.; He, C. H.; Zhang, L. F. Dataset distillation with neural characteristic function: A minmax perspective. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, USA, 2025, pp 25570–25580.
[33] Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, 2015, pp 234–241.
[34] Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; Hochreiter, S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, USA, 2017, pp 6629–6640.
[35] Perez-Cruz, F. Kullback-leibler divergence estimation of continuous distributions. In *Proceedings of the 2008 IEEE International Symposium on Information Theory*, Toronto, Canada, 2008, pp 1666–1670.
[36] van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
[37] Lu, C. X.; Chen, K.; Qiu, H.; Chen, X. J.; Chen, G.; Qi, X. J.; Jiang, H. B. Diffusion-based deep learning method for augmenting ultrastructural imaging and volume electron microscopy. *Nat. Commun.* **2024**, *15*, 4677.
[38] Wang, S.; Liu, X. X.; Li, Y. Y.; Sun, X. Q.; Li, Q.; She, Y. H.; Xu, Y. X.; Huang, X. X.; Lin, R. L.; Kang, D. Y. et al. A deep learning-based stripe self-correction method for stitched microscopic images. *Nat. Commun.* **2023**, *14*, 5393.